

# Effects of Channel State Information Uncertainty on the Performance of Stochastic Signaling

Cagri Goken, Sinan Gezici, and Orhan Arikan

Department of Electrical and Electronics Engineering, Bilkent University, Ankara 06800, Turkey

E-mails: {goken,gezici,oarikan}@ee.bilkent.edu.tr

**Abstract**—In this paper, stochastic signaling is studied for power-constrained scalar valued binary communications systems in the presence of uncertainties in channel state information (CSI). First, it is shown that, for a given decision rule at the receiver, stochastic signaling based on the available CSI at the transmitter results in a randomization between at most two different signal levels for each symbol. Then, the performance of stochastic signaling and conventional deterministic signaling is compared, and sufficient conditions are derived for improvability and nonimprovability of deterministic signaling via stochastic signaling in the presence of CSI uncertainty. Finally a numerical example is presented to explore the theoretical results.

**Index Terms**—Probability of error, stochastic signaling, channel state information.

## I. INTRODUCTION

In binary communications systems over zero-mean additive white Gaussian noise channels and under average power constraints in the form of  $E\{|S_i|^2\} \leq A$  for  $i = 0, 1$ , the average probability of error is minimized when deterministic antipodal signals ( $S_0 = -S_1$ ) are used at the power limit ( $|S_0|^2 = |S_1|^2 = A$ ) and a maximum *a posteriori* probability (MAP) decision rule is used at the receiver [1]. Also, for vector observations, selecting the deterministic signals along the eigenvector of the covariance matrix of the Gaussian noise corresponding to the minimum eigenvalue minimizes the average probability of error [1]. In [2], optimal binary communications over AWGN channels are investigated for nonequal prior probabilities under an average energy per bit constraint, and it is shown that the optimal signaling scheme is on-off keying (OOK) for coherent detection when the signals have nonnegative correlation (also for envelope detection for arbitrary signal correlation).

Although the optimal signaling structures are well-known in the presence of Gaussian noise, the noise can have significantly different probability distribution than the Gaussian distribution in some cases due to effects such as interference and jamming [3]. When the noise is non-Gaussian, *stochastic signaling* can provide performance improvements in terms of average probability of error reduction compared to the conventional deterministic signaling [4]. In the stochastic signaling approach, signals  $S_0$  and  $S_1$  are modeled as random variables whereas they are considered as deterministic quantities in conventional deterministic signaling. For a given decision rule (detector) at the receiver, optimal stochastic signaling is studied under second and fourth moment constraints in [4], and it is shown that an optimal stochastic signal can be represented by a randomization among at most three different signal levels for each symbol. Also, sufficient conditions

are obtained to specify whether stochastic signaling provides improvements over deterministic signaling. In [5], stochastic signaling is studied under an average power constraint in the form of  $\sum_{i=1}^2 \pi_i E\{|S_i|^2\} \leq A$ , where  $\pi_i$  denotes the prior probability of symbol  $i$ , and sufficient conditions are presented to determine performance improvements. In addition, [6] investigates the joint design of the optimal stochastic signals and the detector, and proves that the optimal solution involves randomization between at most two signal levels and the use of the corresponding MAP detector.

Although optimal stochastic signaling for power constrained communications systems has been studied in [4]–[6], no studies have considered the effects of imperfect channel state information (CSI) on the performance of stochastic signaling. In this study, we consider the design of stochastic signals for any given noise probability distribution and detector structure when there are uncertainties in the CSI. After the formulation of the problem, it is shown that an optimal stochastic signal involves randomization between at most two signal levels. Then by deriving upper and lower bounds on the average probability of error for stochastic signaling under CSI uncertainty, sufficient conditions are obtained to specify when the use of stochastic signaling can or cannot improve the performance of conventional signaling. Finally, simulations are performed to illustrate the theoretical results.

## II. SYSTEM MODEL AND MOTIVATION

Consider a binary communications system with scalar observations [7] in which the channel effect can be modeled by a multiplicative term as in flat-fading channels [8], and the received signal is given by

$$Y = \alpha S_i + N, \quad i \in \{0, 1\}, \quad (1)$$

where  $S_0$  and  $S_1$  denote the transmitted signal values for symbol 0 and symbol 1, respectively,  $N$  is the noise component that is independent of  $S_i$ , and  $\alpha$  is the channel coefficient that is modeled as a nonrandom parameter. In addition, the prior probabilities of the symbols, which are denoted by  $\pi_0$  and  $\pi_1$ , are supposed to be known.

In (1), the noise term  $N$  is modeled to have an arbitrary probability distribution considering that it can include the combined effects of thermal noise, interference, and jamming. Hence, the probability distribution of the noise component is not necessarily Gaussian [3], [7].

A generic decision rule is considered at the receiver to determine the symbol in (1). For a given observation  $Y = y$ , the decision rule  $\phi(y)$  is expressed as

$$\phi(y) = \begin{cases} 0, & y \in \Gamma_0 \\ 1, & y \in \Gamma_1 \end{cases}, \quad (2)$$

<sup>0</sup>This research was supported in part by the National Young Researchers Career Development Programme (project no. 110E245) of the Scientific and Technological Research Council of Turkey (TUBITAK).

where  $\Gamma_0$  and  $\Gamma_1$  are the decision regions for symbol 0 and symbol 1, respectively [1].

The aim is to design signals  $S_0$  and  $S_1$  in (1) in order to minimize the average probability of error for a given decision rule, which is given by

$$P_{\text{avg}} = \pi_0 P_0(\Gamma_1) + \pi_1 P_1(\Gamma_0), \quad (3)$$

with  $P_i(\Gamma_j)$  denoting the probability of selecting symbol  $j$  when symbol  $i$  is transmitted. In practical systems, there exists an average power constraint on the signals, which can be expressed as

$$E\{|S_i|^2\} \leq A, \quad (4)$$

for  $i = 0, 1$ , where  $A$  is the average power limit. Therefore, in the stochastic signaling approach, the aim becomes the calculation of the optimal probability density functions (PDFs) for signals  $S_0$  and  $S_1$  that minimize the average probability of error in (3) under the average power constraint in (4) [4].

Unlike stochastic signaling, in the conventional signal design,  $S_0$  and  $S_1$  are modeled as deterministic signals and set to  $S_0 = -\sqrt{A}$  and  $S_1 = \sqrt{A}$  [1], [8]. Then, the average probability of error in (3) becomes

$$P_{\text{conv}} = \pi_0 \int_{\Gamma_1} p_N(y + \alpha\sqrt{A}) dy + \pi_1 \int_{\Gamma_0} p_N(y - \alpha\sqrt{A}) dy, \quad (5)$$

where  $p_N(\cdot)$  is the PDF of the noise in (1).

As investigated in [4]-[6], stochastic signaling results in lower average probabilities of error than the conventional deterministic signaling in some cases in the presence of non-Gaussian noise. However, the common assumption in the previous studies is that the channel coefficient  $\alpha$  in (1) is known perfectly at the transmitter, i.e., the CSI is available at the transmitter. In practice, the transmitter can obtain CSI via feedback from the receiver, or by utilizing the reciprocity of forward and reverse links under time division duplexing [9]. In both scenarios, it is realistic to model the CSI at the transmitter to include certain errors/uncertainties. Therefore, the main motivation behind this study is to investigate stochastic signaling under imperfect CSI; that is, to evaluate the performance of stochastic signaling in practical scenarios.

### III. OPTIMAL STOCHASTIC SIGNALING

Let  $p_{S_0}(\cdot)$  and  $p_{S_1}(\cdot)$  denote the PDFs of  $S_0$  and  $S_1$  in (1), respectively. Also define  $\hat{S}_0 \triangleq \alpha S_0$  and  $\hat{S}_1 \triangleq \alpha S_1$ , and denote their PDFs as  $p_{\hat{S}_0}(\cdot)$  and  $p_{\hat{S}_1}(\cdot)$ , respectively. Then, from (3), the average probability of error for the decision rule in (2) is given by

$$P_{\text{stoc}} = \sum_{i=0}^1 \pi_i \int_{-\infty}^{\infty} p_{\hat{S}_i}(t) \int_{\Gamma_{1-i}} p_N(y - t) dy dt. \quad (6)$$

Since  $p_{\hat{S}_i}(t)$  is given by  $p_{\hat{S}_i}(t) = (1/|\alpha|) p_{S_i}(t/\alpha)$  for  $i = 0, 1$ , (6) can be expressed, after a change of variable ( $t = \alpha x$ ), as

$$P_{\text{stoc}} = \sum_{i=0}^1 \pi_i \int_{-\infty}^{\infty} p_{S_i}(x) \int_{\Gamma_{1-i}} p_N(y - \alpha x) dy dx. \quad (7)$$

Since imperfect CSI is considered in this study, the transmitter has a distorted version of the correct channel coefficient

$\alpha$ . Let  $\hat{\alpha}$  denote this distorted (noisy) channel coefficient at the transmitter. The transmitter employs  $\hat{\alpha}$  in the design of stochastic signals. Then, the stochastic signal design problem can be expressed as

$$\begin{aligned} \min_{p_{S_0}, p_{S_1}} \quad & \sum_{i=0}^1 \pi_i \int_{-\infty}^{\infty} p_{S_i}(x) \int_{\Gamma_{1-i}} p_N(y - \hat{\alpha} x) dy dx \\ \text{subject to} \quad & E\{|S_i|^2\} \leq A, \quad i = 0, 1. \end{aligned} \quad (8)$$

Note that there are also implicit constraints in the optimization problem in (8) because  $p_{S_0}(\cdot)$  and  $p_{S_1}(\cdot)$  need to satisfy the conditions to be valid PDFs. As in [4], this optimization problem can be expressed as two separate optimization problems for  $S_0$  and  $S_1$ . Namely, the optimal signal PDF for symbol 1 can be obtained from the solution of the following optimization problem:

$$\begin{aligned} \min_{p_{S_1}} \quad & \int_{-\infty}^{\infty} p_{S_1}(x) \int_{\Gamma_0} p_N(y - \hat{\alpha} x) dy dx \\ \text{subject to} \quad & E\{|S_1|^2\} \leq A. \end{aligned} \quad (9)$$

If  $G(x, k)$  is defined as

$$G(x, k) \triangleq \int_{\Gamma_0} p_N(y - kx) dy, \quad (10)$$

(9) can also be written as

$$\min_{p_{S_1}} E\{G(S_1, \hat{\alpha})\} \quad \text{subject to} \quad E\{|S_1|^2\} \leq A, \quad (11)$$

where the expectations are taken over  $S_1$ . Note that  $G(S_1, \hat{\alpha})$  is only a function of  $S_1$  for a given fixed  $\hat{\alpha}$ . In the previous studies, such as [4] and [10], the optimization problems with the same structure as (11) have been explored thoroughly. If  $G(S_1, \hat{\alpha})$  in (11) is a continuous function of  $S_1$ , and  $S_1$  takes values in  $[-\gamma, \gamma]$  for some finite  $\gamma > 0$ , then the optimal solution of (11) can be represented by a randomization of at most two signal levels as a result of Carathéodory's theorem [11]. Hence, the optimal signal PDF for  $S_1$  can be expressed as

$$p_{S_1}(s) = \lambda_1 \delta(s - s_{11}) + (1 - \lambda_1) \delta(s - s_{12}), \quad (12)$$

where  $\lambda_1 \in [0, 1]$ .

A similar optimization problem can also be formulated for  $S_0$ . After obtaining the optimal signal PDFs for  $S_0$  and  $S_1$ , the corresponding average probability of error can be calculated. Since the optimization problems are similar for  $S_0$  and  $S_1$ , we focus on the design of  $S_1$  in the remainder of this study.

### IV. STOCHASTIC SIGNALING VERSUS CONVENTIONAL SIGNALING

It is known that, in the presence of perfect CSI at the transmitter, conventional signaling, which sets  $S_1 = \sqrt{A}$  [that is,  $p_{S_1}(x) = \delta(x - \sqrt{A})$ ], can or cannot be optimal under certain sufficient conditions as discussed in [4]. In this section, we explore the conditions under which the use of stochastic signaling instead of deterministic signaling can result in improved average probability of error performance in the presence of *imperfect* CSI.

In the presence of imperfect CSI, let the transmitter have the channel coefficient information as  $\hat{\alpha}$ . Then, the transmitter obtains the optimal stochastic signal  $S_1$  from (11). Let  $p_{S_1}^{\hat{\alpha}}(\cdot)$  denote the solution of (11) for a given value of  $\hat{\alpha}$ . Then, the

corresponding conditional probability of error for symbol 1 can be expressed as

$$P_e^{\hat{\alpha}} = \int_{-\infty}^{\infty} p_{S_1}^{\hat{\alpha}}(x) G(x, \alpha) dx, \quad (13)$$

where  $G(x, \alpha)$  is as defined in (10). Note that  $G(x, \alpha)$  specifies the probability of choosing symbol 0 for a given signal value  $x$  for symbol 1 when the channel coefficient is equal to  $\alpha$ . Therefore, when the stochastic signal for symbol 1 is specified by the PDF  $p_{S_1}^{\hat{\alpha}}(x)$ , the corresponding conditional probability of error for symbol 1 is obtained as in (13).

Suppose that  $\hat{\alpha}$  can be modeled as a random variable with a generic PDF  $p_{\hat{\alpha}}(\cdot)$ . In order to improve the performance of conventional signaling for symbol 1 via stochastic signaling, we need to have  $P_e < G(\sqrt{A}, \alpha)$ , where  $G(\sqrt{A}, \alpha)$  is the conditional probability of error for conventional signaling, i.e., for  $S_1 = \sqrt{A}$  (see (5) and (10)), and  $P_e$  is the average conditional probability of error for stochastic signaling based on imperfect CSI, which can be calculated as

$$P_e = \int_{-\infty}^{\infty} p_{\hat{\alpha}}(a) P_e^a da, \quad (14)$$

with  $P_e^a$  being given by (13).

In order to derive sufficient conditions for the improbability and nonimprovability of conventional signaling via stochastic signaling, assume that the channel coefficient information at the transmitter is specified as  $\hat{\alpha} = \alpha + \eta$ , where  $\eta$  is a zero-mean Gaussian noise with standard deviation  $\varepsilon$ ; that is,  $\eta \sim \mathcal{N}(0, \varepsilon^2)$ . Although the Gaussian error model is employed for the convenience of the analysis, the results are valid also for non-Gaussian error models, as will be discussed at the end of this section. In addition, it is assumed that  $\alpha$  is a positive number without loss of generality.<sup>1</sup> Then, the following proposition presents sufficient conditions on the improbability and nonimprovability of conventional signaling via stochastic signaling.

**Proposition 1:** Assume that  $G(x, k)$  in (10) and  $P_e^{\hat{\alpha}}$  in (13) have the following properties:

- $G(x, k)$  is a strictly decreasing function of  $x$  for any fixed positive  $k$ , and  $G(x, k) = 1 - G(-x, k)$ .
- There exist  $\kappa_1, \kappa_2, \gamma_{th}, \theta_{th}$ , and  $\beta_{th}$  such that  $P_e^{\hat{\alpha}} < \kappa_1$  when  $\hat{\alpha} > \gamma_{th} > 0$ ;  $P_e^{\hat{\alpha}} < \kappa_2 < \kappa_1$  when  $\alpha > \hat{\alpha} > \theta_{th} > \gamma_{th}$ ; and  $P_e^{\hat{\alpha}} = G(\sqrt{A}, \alpha)$  when  $\hat{\alpha} > \beta_{th} > \alpha$ .

Then, stochastic signaling performs worse than conventional signaling if the standard deviation  $\varepsilon$  of the channel coefficient error satisfies following inequality

$$\begin{aligned} & \left( \frac{1}{2} - \kappa_1 \right) Q \left( \frac{\alpha + \gamma_{th}}{\varepsilon} \right) + \frac{1}{2} Q \left( \frac{\alpha}{\varepsilon} \right) \\ & + (\kappa_1 - \kappa_2) \left( Q \left( \frac{2\alpha}{\varepsilon} \right) - Q \left( \frac{\alpha + \theta_{th}}{\varepsilon} \right) \right) \\ & + Q \left( \frac{\beta_{th} - \alpha}{\varepsilon} \right) G(\sqrt{A}, \alpha) \geq G(\sqrt{A}, \alpha), \end{aligned} \quad (15)$$

and stochastic signaling performs better than conventional signaling if  $\varepsilon$  satisfies following inequality<sup>2</sup>

<sup>1</sup>If it is negative, one can redefine the  $G$  function in (10) by using  $p_N(y + kx)$  instead of  $p_N(y - kx)$ .

<sup>2</sup>Note that the choice of parameters in the conditions of Proposition 1 is important to satisfy the inequalities in (15) and (16). Also, the  $Q$ -function is defined as  $Q(x) = (\int_x^{\infty} e^{-t^2/2} dt) / \sqrt{2\pi}$ .

$$\begin{aligned} & \frac{1}{2} \left( \kappa_1 + \kappa_2 + Q \left( \frac{\alpha}{\varepsilon} \right) \right) + \left( \frac{1}{2} - \kappa_1 \right) Q \left( \frac{\alpha - \gamma_{th}}{\varepsilon} \right) \\ & - \kappa_1 Q \left( \frac{\beta_{th} - \alpha}{\varepsilon} \right) + (\kappa_1 - \kappa_2) Q \left( \frac{\alpha - \theta_{th}}{\varepsilon} \right) + \\ & \left( Q \left( \frac{\beta_{th} - \alpha}{\varepsilon} \right) - Q \left( \frac{\alpha + \beta_{th}}{\varepsilon} \right) \right) G(\sqrt{A}, \alpha) \leq G(\sqrt{A}, \alpha). \end{aligned} \quad (16)$$

**Proof:** Please see Appendix A.

Although the results in Proposition 1 are presented for channel coefficient errors with a zero-mean Gaussian distribution, they can easily be extended for any type of probability distribution as well. For example, consider a generic PDF for the channel coefficient error, which is denoted by  $p_{\eta}(\cdot)$ . The corresponding cumulative distribution function (CDF)  $F_{\eta}(\cdot)$  can be expressed as  $F_{\eta}(x) = \int_{-\infty}^x p_{\eta}(t) dt$ . Then, the results in Proposition 1 are still valid when  $Q(x/\varepsilon)$  in (15) and (16) are replaced by  $1 - F_{\eta}(x)$ .

As discussed before,  $G(x, k)$  can be inferred as the probability of deciding symbol 0 instead of symbol 1, when the value of the channel coefficient is  $k$ , and  $S_1 = x$ . In general, for a specific channel coefficient, when a larger signal value is employed, a lower probability of error can be obtained; hence,  $G(x, k)$  is usually a decreasing function of  $x$  in practice. Moreover,  $G(x, k) = 1 - G(-x, k)$  can be satisfied when the channel noise has a symmetric PDF (i.e.  $p_N(x) = p_N(-x)$ ) and the decision regions of the detector at the receiver are symmetric ( $\Gamma_0 = -\Gamma_1$ ). In fact, the channel noise is symmetric in most practical scenarios (for example, zero-mean additive Gaussian noise or Gaussian mixture noise with symmetric components [3]) and some receivers such as the sign detector or the optimal MAP detector for symmetric signaling under symmetric channel noise will have symmetric decision regions in fact. All in all, the first condition in the proposition is expected to hold in many practical scenarios. The details of how the second condition is satisfied and the parameters are chosen will be investigated in the next section.

## V. PERFORMANCE EVALUATION

In this section, a numerical example is presented to compare the performance of conventional signaling and stochastic signaling in the presence of channel coefficient errors. A binary communications system with equally likely symbols are considered ( $\pi_0 = \pi_1 = 0.5$ ). In addition, the average power limit in (4) is set to  $A = 1$ . We assume that the decision rule at the receiver is specified by  $\Gamma_0 = (-\infty, 0]$  and  $\Gamma_1 = [0, \infty)$  (i.e., the sign detector). The noise in (1) is modeled by a Gaussian mixture noise [3] with its PDF being given by

$$p_N(n) = \frac{1}{\sqrt{2\pi} \sigma} \sum_{l=1}^L v_l e^{-\frac{(n-\mu_l)^2}{2\sigma^2}}. \quad (17)$$

Gaussian mixture noise is encountered in practical systems in the presence of interference [3]. For the channel noise and detector structure as described above  $G(x, k)$  in (10) can be calculated as

$$G(x, k) = \sum_{l=1}^L v_l Q \left( \frac{kx + \mu_l}{\sigma} \right). \quad (18)$$



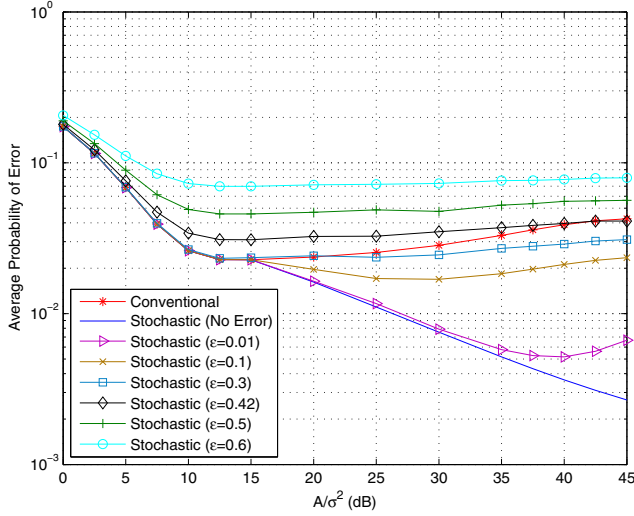


Fig. 1. Average probability of error versus  $A/\sigma^2$  for conventional signaling and stochastic signaling with various  $\varepsilon$  values.

In this example, the mass points  $\mu_l$  are located at  $\boldsymbol{\mu} = [-1.013 \ -0.275 \ -0.105 \ 0.105 \ 0.275 \ 1.013]$  with corresponding weights  $\mathbf{v} = [0.043 \ 0.328 \ 0.129 \ 0.129 \ 0.328 \ 0.043]$ . Note that each component of the Gaussian mixture noise has the same variance  $\sigma^2$  and the average power of the noise can be calculated as  $E\{n^2\} = \sigma^2 + 0.1407$ .

The channel coefficient information at the transmitter is modeled as  $\hat{\alpha} = \alpha + \eta$ , where  $\alpha = 1$  and  $\eta$  is a zero-mean Gaussian random variable with standard deviation  $\varepsilon$ . Due to the symmetry of the problem, the conditional probability of error expression in (14) also provides the average probability of error in this scenario. In order to evaluate that expression, 100 realizations are obtained for  $\hat{\alpha}$ . Then, the optimization problem in (11) is solved for each realization and the optimal signal PDFs that are in the form of (12) are obtained by using the particle swarm optimization (PSO) approach [12]. For the details of the PSO parameters employed in this study, please refer to [6].

In Fig. 1, the average probabilities of error are plotted versus  $A/\sigma^2$  for conventional signaling, stochastic signaling with no channel coefficient errors ( $\varepsilon = 0$ ), and stochastic signaling with various levels of channel coefficient errors. It is observed for high  $A/\sigma^2$  values that the best performance is obtained by stochastic signaling with perfect CSI and the performance of stochastic signaling gets worse as the variance of the channel coefficient error increases. For example, when  $\varepsilon = 0.5$  and  $\varepsilon = 0.6$ , stochastic signaling performs worse than conventional signaling for all  $A/\sigma^2$  values. Another observation is that for low values of  $\varepsilon$ , stochastic signaling still performs better than conventional signaling for high  $A/\sigma^2$  values and their performance is similar for high  $\sigma^2$ , i.e. when  $A/\sigma^2$  is smaller than 15 dB. In fact, one can calculate the average probability of error analytically for low  $A/\sigma^2$  values for each  $\varepsilon$ . At low  $A/\sigma^2$  values,  $P_e^{\hat{\alpha}}$  in (13) can be expressed as

$$P_e^{\hat{\alpha}} = \frac{1 - \text{sgn}(\hat{\alpha})}{2} + \text{sgn}(\hat{\alpha}) G(\sqrt{A}, \alpha), \quad (19)$$

where  $\text{sgn}$  denotes the sign operator. Then, from (14),  $P_e$  can be calculated as  $Q(\alpha/\varepsilon) + G(\sqrt{A}, \alpha) - 2G(\sqrt{A}, \alpha)Q(\alpha/\varepsilon)$ .

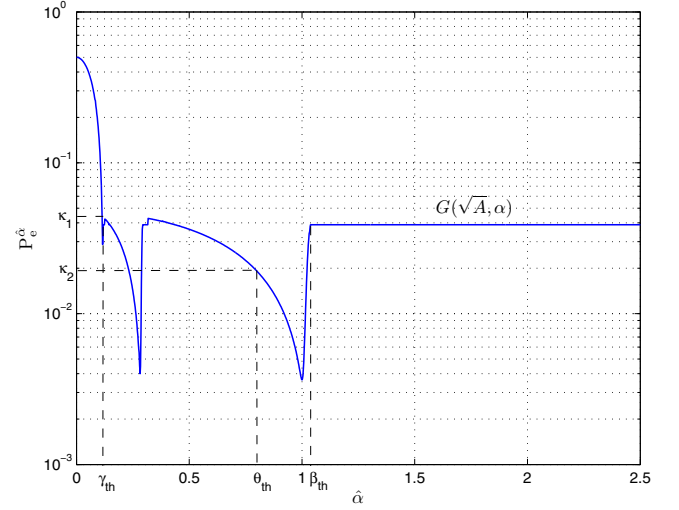


Fig. 2.  $P_e^{\hat{\alpha}}$  versus  $\hat{\alpha}$  for  $A/\sigma^2 = 40$  dB. The second condition in Proposition 1 is satisfied for  $\kappa_1 = 0.04354$ ,  $\kappa_2 = 0.01913$ ,  $\gamma_{th} = 0.1135$ ,  $\theta_{th} = 0.8$ ,  $\beta_{th} = 1.038$ , and  $G(\sqrt{A}, \alpha) = 0.03884$ .

For instance, when  $A/\sigma^2 = 10$  dB,  $G(\sqrt{A}, \alpha) = 0.02613$  in this example. Then, for  $\varepsilon = 0.6$ ,  $P_e$  is calculated as  $0.9477 Q(5/3) + 0.0261 = 0.0714$ , which matches with the result shown in Fig. 1. For this example, we can apply the conditions given in Proposition 1 and check if the conventional signaling is improvable or nonimprovable via stochastic signaling for given  $\varepsilon$  values. Firstly, we examine the first condition in the proposition.  $G(x, k)$  is calculated above for this example and it is a convex combination of  $Q$  functions. Therefore,  $G(x, k)$  is a strictly decreasing function of  $x$  as  $Q(x)$  is a monotone decreasing function. Also, since  $Q(x) = 1 - Q(-x)$  and the components of Gaussian mixture noise are symmetric, we have  $G(x, k) = 1 - G(-x, k)$  as well. Hence, the first condition in Proposition 1 is satisfied. In order to check the second condition, the plot of  $P_e^{\hat{\alpha}}$  versus  $\hat{\alpha}$  is presented in Fig. 2 for  $A/\sigma^2 = 40$  dB. It is observed that  $P_e^{\hat{\alpha}}$  does not have a monotonic structure; that is, it increases, decreases or remains the same as  $\hat{\alpha}$  increases. However, it obeys the structure specified in the second condition of Proposition 1. Specifically, when  $\hat{\alpha} > \gamma_{th} = 0.1135$ ,  $P_e^{\hat{\alpha}}$  is less than  $\kappa_1 = 0.04354$ , and when  $\theta_{th} = 0.8 < \hat{\alpha} < \alpha = 1$ ,  $P_e^{\hat{\alpha}}$  becomes less than  $\kappa_2 = 0.01913$ , which is even smaller than  $\kappa_1$ . Also, when  $\hat{\alpha} > \beta_{th} = 1.038$ ,  $P_e^{\hat{\alpha}}$  becomes equal to  $G(\sqrt{A}, \alpha) = 0.03884$ , which is the average probability of error for conventional signaling. The values of  $\kappa_1$ ,  $\kappa_2$ ,  $\gamma_{th}$ ,  $\theta_{th}$ , and  $\beta_{th}$  are illustrated in Fig. 2. Based on the specified parameters, (15) becomes

$$0.45646 Q\left(\frac{1.1135}{\varepsilon}\right) + 0.02441 \left( Q\left(\frac{2}{\varepsilon}\right) - Q\left(\frac{1.8}{\varepsilon}\right) \right) + 0.5 Q\left(\frac{1}{\varepsilon}\right) + 0.03884 Q\left(\frac{0.038}{\varepsilon}\right) \geq 0.03884.$$

For  $\varepsilon = 0.6$ , the left-hand-side of this inequality is calculated to be 0.0568; hence, the inequality is satisfied. This means that when  $A/\sigma^2 = 40$  dB, if the standard deviation of the channel coefficient error equals to 0.6, we can conclude that stochastic signaling is outperformed by conventional signaling. In fact,

it can be observed from Fig.1 that for  $A/\sigma^2 = 40$  dB and  $\varepsilon = 0.6$ , the performance of stochastic signaling is quite worse than that of conventional signaling as Proposition 1 asserts. Also note that when  $\varepsilon^* = 0.5178$ , (15) becomes an equality.

Similarly, based on the selected parameters, (16) becomes

$$\begin{aligned} & 0.5 \left( 0.06267 + Q\left(\frac{1}{\varepsilon}\right) \right) + 0.45646 Q\left(\frac{0.8865}{\varepsilon}\right) \\ & - 0.04354 Q\left(\frac{0.038}{\varepsilon}\right) + 0.02441 Q\left(\frac{0.2}{\varepsilon}\right) \\ & + 0.03884 \left( Q\left(\frac{0.038}{\varepsilon}\right) - Q\left(\frac{2.038}{\varepsilon}\right) \right) \leq 0.03884. \end{aligned}$$

For  $\varepsilon = 0.3, 0.1, 0.01$ , the left-hand-side of the inequality is calculated as 0.03631, 0.03024, 0.03133, respectively. This means that, at  $A/\sigma^2 = 40$  dB, if the standard deviation of the channel coefficient error is 0.3, 0.1, or 0.01, we can conclude that conventional signaling is outperformed by stochastic signaling as a result of Proposition 1. This can also be observed from Fig. 1 when  $A/\sigma^2 = 40$  dB for  $\varepsilon = 0.3, 0.1, 0.01$ . Also, when  $\hat{\varepsilon} = 0.3395$ , (16) turns out to be an equality.

In order to explore the performance of stochastic signaling with respect to  $\varepsilon$ , Fig. 3 is presented. It is observed that as the variance of the channel coefficient error increases, the average probability of error for stochastic signaling increases as well. This is expected since the transmitter designs the stochastic signals in the presence of channel coefficient errors (imperfect CSI) and these errors get more significant as  $\varepsilon$  increases. Therefore, it can be concluded that in the presence of large channel coefficient errors (i.e., large  $\varepsilon$ ), using conventional deterministic signaling instead of stochastic signaling can be more preferable, whereas for small channel coefficient errors, stochastic signaling can be employed to achieve smaller average probabilities of error than conventional signaling. In Fig. 3,  $\varepsilon^*$  and  $\hat{\varepsilon}$  are also illustrated, together with the point  $\varepsilon_{th}$  at which the performance of stochastic signaling and conventional signaling becomes the same. Note that the conditions in Proposition 1 are not necessary but only sufficient conditions for the improvability and nonimprovability of conventional signal via stochastic signaling.

## VI. CONCLUSIONS

Stochastic signaling has been investigated in the presence of imperfect CSI. First, a problem formulation has been presented in the presence of errors in the channel coefficient, and the two mass point structure of an optimal signal PDF has been mentioned. Then, upper and lower bounds on the average probability of error of stochastic signaling under imperfect CSI have been derived and sufficient conditions have been presented to specify when the performance of conventional deterministic signaling can or cannot be improved via stochastic signaling. The theoretical results have been explained over an example.

## REFERENCES

- [1] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
- [2] I. Korn, J. P. Fonseka, and S. Xing, "Optimal binary communication with nonequal probabilities," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1435–1438, Sep. 2003.
- [3] V. Bhatia and B. Mulgrew, "Non-parametric likelihood based channel estimator for Gaussian mixture noise," *Signal Processing*, vol. 87, pp. 2569–2586, Nov. 2007.

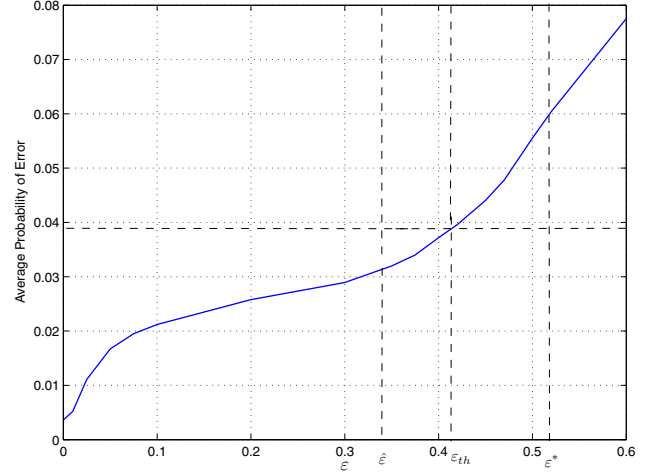


Fig. 3. Average probability of error versus  $\varepsilon$  for stochastic signaling. At  $\varepsilon_{th} = 0.413$ , stochastic signaling has the same average probability of error as conventional signaling.

- [4] C. Goken, S. Gezici, and O. Arikan, "Optimal stochastic signaling for power-constrained binary communications systems," *IEEE Trans. on Wireless Commun.*, vol. 9, no. 12, pp. 3650–3661, Dec. 2010.
- [5] —, "On the optimality of stochastic signaling under an average power constraint," in *Proc. 48th Annual Allerton Conf. on Commun., Control, and Computing*, Illinois, Sep. 29-Oct. 1 2010, pp. 1158–1164.
- [6] —, "Optimal signaling and detector design for power-constrained binary communications systems over non-Gaussian channels," *IEEE Commun. Lett.*, vol. 14, no. 2, pp. 100–102, Feb. 2010.
- [7] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inform. Theory*, vol. 42, no. 4, pp. 1316–1321, July 1996.
- [8] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
- [9] A. Goldsmith, *Wireless Communications*. Cambridge, UK: Cambridge University Press, 2005.
- [10] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection," *IEEE Trans. Sig. Processing*, vol. 57, no. 5, pp. 1655–1669, May 2009.
- [11] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Berlin: Springer-Verlag, 2004.
- [12] A. I. F. Vaz and E. M. G. P. Fernandes, "Optimization of nonlinear constrained particle swarm," *Baltic Journal on Sustainability*, vol. 12, no. 1, pp. 30–36, 2006.

## APPENDIX

### A. Proof of Proposition 1

In the following, lower and upper bounds for the expression in (14) are derived in order to prove the statements in the proposition. We start by noticing the fact that the sign of the channel coefficient knowledge at the transmitter is important. Suppose that  $p_{S_1}^{\hat{\alpha}}$  is the optimal PDF obtained from (11) for a given  $\hat{\alpha}$ . Therefore, if  $-\hat{\alpha}$  is used instead of  $\hat{\alpha}$ , then  $p_{S_1}^{-\hat{\alpha}}$  will be the optimal solution of (11) and the value of  $p_{S_1}^{-\hat{\alpha}}(x)$  will be equal to  $p_{S_1}^{\hat{\alpha}}(-x)$ . This observation can be utilized in (13), and also using the fact that  $G(x, k) = 1 - G(-x, k)$ ,  $P_e^{\hat{\alpha}} = 1 - P_e^{-\hat{\alpha}}$  can be obtained as follows:

$$\begin{aligned} & \int_{-\infty}^{\infty} p_{S_1}^{\hat{\alpha}}(x) G(x, k) dx = \int_{-\infty}^{\infty} p_{S_1}^{-\hat{\alpha}}(-x) (1 - G(-x, k)) dx \\ & = \int_{-\infty}^{\infty} p_{S_1}^{-\hat{\alpha}}(t) (1 - G(t, k)) dt = 1 - \int_{-\infty}^{\infty} p_{S_1}^{-\hat{\alpha}}(t) G(t, k) dt \\ & = 1 - P_e^{-\hat{\alpha}}. \end{aligned} \quad (20)$$

It is stated in the second condition of the proposition that  $P_e^{\hat{\alpha}} < \kappa_1$  when  $\hat{\alpha} > \gamma_{th}$ , and  $P_e^{\hat{\alpha}} < \kappa_2 < \kappa_1$  when  $\alpha > \hat{\alpha} > \theta_{th}$ . Therefore, if we insert  $-\hat{\alpha}$  instead of  $\hat{\alpha}$  in these conditions, we get  $P_e^{-\hat{\alpha}} < \kappa_1$  when  $-\hat{\alpha} > \gamma_{th}$  and  $P_e^{-\hat{\alpha}} < \kappa_2 < \kappa_1$  when  $\alpha > -\hat{\alpha} > \theta_{th}$ . Using the result in (20) and rearranging the terms yield  $P_e^{\hat{\alpha}} > 1 - \kappa_1$  when  $\hat{\alpha} < -\gamma_{th}$  and  $P_e^{\hat{\alpha}} > 1 - \kappa_2 > 1 - \kappa_1$  when  $-\alpha < \hat{\alpha} < -\theta_{th}$ . Also, since  $G(x, k)$  is a strictly decreasing function of  $x$  when  $k$  is positive, then  $G(x, \hat{\alpha})$  is a strictly increasing function of  $x$  if  $\hat{\alpha}$  is negative. Therefore, for a given  $\hat{\alpha} < 0$ , the optimal signal PDF  $p_{S_1}^{\hat{\alpha}}$  assigns the weights on negative numbers instead of positive ones since for each positive value of  $S_1$ , its negative can be used instead, which results in the same average power value and a smaller  $E\{G(S_1, \hat{\alpha})\}$ . Furthermore, since  $G(x, \alpha)$  is a strictly decreasing function, and  $G(x, \alpha) = 1 - G(-x, \alpha)$ , we have  $G(x, \alpha) > G(0, \alpha) = 0.5$  for  $x < 0$ . Thus, by using these two facts and the expression in (13), we conclude that if  $\hat{\alpha} < 0$ , then  $P_e^{\hat{\alpha}} > 0.5$  [and  $P_e^{\hat{\alpha}} < 0.5$ , if  $\hat{\alpha} > 0$ ]. Now, one can find a lower bound on  $P_e$  in (14) as follows:

$$\begin{aligned}
 P_e &= \int_{-\infty}^{\infty} p_{\hat{\alpha}}(a) P_e^a da \\
 &\geq \int_{-\infty}^{-\gamma_{th}} p_{\hat{\alpha}}(a) P_e^a da + \int_{-\gamma_{th}}^0 p_{\hat{\alpha}}(a) P_e^a da + \int_{\beta_{th}}^{\infty} p_{\hat{\alpha}}(a) P_e^a da \\
 &> (1 - \kappa_1) P(\hat{\alpha} < -\gamma_{th}) + (\kappa_1 - \kappa_2) P(-\alpha < \hat{\alpha} < -\theta_{th}) \\
 &+ \frac{1}{2} P(-\gamma_{th} < \hat{\alpha} < 0) + P(\hat{\alpha} > \beta_{th}) G(\sqrt{A}, \alpha) \\
 &= (1 - \kappa_1) P\left(\frac{\eta}{\varepsilon} > \frac{\alpha + \gamma_{th}}{\varepsilon}\right) \\
 &+ (\kappa_1 - \kappa_2) P\left(\frac{-2\alpha}{\varepsilon} < \frac{\eta}{\varepsilon} < \frac{-\alpha - \theta_{th}}{\varepsilon}\right) \\
 &+ \frac{1}{2} P\left(\frac{-\alpha}{\varepsilon} < \frac{\eta}{\varepsilon} < \frac{-\alpha - \gamma_{th}}{\varepsilon}\right) \\
 &+ P\left(\frac{\eta}{\varepsilon} > \frac{\beta_{th} - \alpha}{\varepsilon}\right) G(\sqrt{A}, \alpha) \\
 &= (1 - \kappa_1) Q\left(\frac{\alpha + \gamma_{th}}{\varepsilon}\right) \\
 &+ (\kappa_1 - \kappa_2) \left(Q\left(\frac{2\alpha}{\varepsilon}\right) - Q\left(\frac{\alpha + \theta_{th}}{\varepsilon}\right)\right) \\
 &+ \frac{1}{2} \left(Q\left(\frac{\alpha}{\varepsilon}\right) - Q\left(\frac{\alpha + \gamma_{th}}{\varepsilon}\right)\right) + Q\left(\frac{\beta_{th} - \alpha}{\varepsilon}\right) G(\sqrt{A}, \alpha) \\
 &= \left(\frac{1}{2} - \kappa_1\right) Q\left(\frac{\alpha + \gamma_{th}}{\varepsilon}\right) \\
 &+ (\kappa_1 - \kappa_2) \left(Q\left(\frac{2\alpha}{\varepsilon}\right) - Q\left(\frac{\alpha + \theta_{th}}{\varepsilon}\right)\right) \\
 &+ \frac{1}{2} Q\left(\frac{\alpha}{\varepsilon}\right) + Q\left(\frac{\beta_{th} - \alpha}{\varepsilon}\right) G(\sqrt{A}, \alpha). \quad (21)
 \end{aligned}$$

Note that the first inequality follows from the fact that a positive term, namely,  $\int_0^{\beta_{th}} p_{\hat{\alpha}}(a) P_e^a da$ , is removed from the initial expression  $\int_{-\infty}^{\infty} p_{\hat{\alpha}}(a) P_e^a da$ . Also, in obtaining the first and the second terms after the second inequality, we use the fact that although  $P_e^{\hat{\alpha}} > 1 - \kappa_1$  when  $\hat{\alpha} < -\gamma_{th}$ , the bound is tighter, that is,  $P_e^{\hat{\alpha}} > 1 - \kappa_2$ , when  $-\alpha < \hat{\alpha} < -\theta_{th} < -\gamma_{th}$ .

For a given  $\varepsilon$ , if the final expression in (21) is greater than or equal to  $G(\sqrt{A}, \alpha)$ , then  $P_e > G(\sqrt{A}, \alpha)$ . Therefore,

under the conditions in the proposition, if the inequality in (15) is satisfied for a given value of the standard deviation  $\varepsilon$  of the channel coefficient error, it is sufficient to conclude that conventional signaling performs better than stochastic signaling.

Next, one can find an upper bound on  $P_e$  in (14) as follows:

$$\begin{aligned}
 P_e &= \int_{-\infty}^{\infty} p_{\hat{\alpha}}(a) P_e^a da \\
 &= \int_{-\infty}^{-\beta_{th}} p_{\hat{\alpha}}(a) P_e^a da + \int_{-\beta_{th}}^0 p_{\hat{\alpha}}(a) P_e^a da \\
 &+ \int_0^{\gamma_{th}} p_{\hat{\alpha}}(a) P_e^a da + \int_{\gamma_{th}}^{\beta_{th}} p_{\hat{\alpha}}(a) P_e^a da + \int_{\beta_{th}}^{\infty} p_{\hat{\alpha}}(a) P_e^a da \\
 &< (1 - G(\sqrt{A}, \alpha)) P(\hat{\alpha} < -\beta_{th}) + P(-\beta_{th} < \hat{\alpha} < 0) \\
 &+ \frac{1}{2} P(0 < \hat{\alpha} < \gamma_{th}) + \kappa_1 P(\gamma_{th} < \hat{\alpha} < \beta_{th}) \\
 &+ (\kappa_2 - \kappa_1) P(\theta_{th} < \hat{\alpha} < \alpha) + P(\hat{\alpha} > \beta_{th}) G(\sqrt{A}, \alpha) \\
 &= (1 - G(\sqrt{A}, \alpha)) P\left(\frac{\eta}{\varepsilon} > \frac{\alpha + \beta_{th}}{\varepsilon}\right) \\
 &+ P\left(\frac{-\beta_{th} - \alpha}{\varepsilon} < \frac{\eta}{\varepsilon} < \frac{-\alpha}{\varepsilon}\right) \\
 &+ \frac{1}{2} P\left(\frac{-\alpha}{\varepsilon} < \frac{\eta}{\varepsilon} < \frac{-\alpha + \gamma_{th}}{\varepsilon}\right) + \kappa_1 P\left(\frac{-\alpha + \gamma_{th}}{\varepsilon} < \frac{\eta}{\varepsilon} < \frac{-\alpha + \beta_{th}}{\varepsilon}\right) \\
 &+ (\kappa_2 - \kappa_1) P\left(\frac{-\alpha + \theta_{th}}{\varepsilon} < \frac{\eta}{\varepsilon} < 0\right) \\
 &+ P\left(\frac{\eta}{\varepsilon} > \frac{\beta_{th} - \alpha}{\varepsilon}\right) G(\sqrt{A}, \alpha) \\
 &= (1 - G(\sqrt{A}, \alpha)) Q\left(\frac{\alpha + \beta_{th}}{\varepsilon}\right) + \frac{1}{2} \left(Q\left(\frac{\alpha - \gamma_{th}}{\varepsilon}\right) - Q\left(\frac{\alpha}{\varepsilon}\right)\right) \\
 &+ Q\left(\frac{\alpha}{\varepsilon}\right) - Q\left(\frac{\alpha + \beta_{th}}{\varepsilon}\right) \\
 &+ \kappa_1 \left(1 - Q\left(\frac{\alpha - \gamma_{th}}{\varepsilon}\right) - Q\left(\frac{\beta_{th} - \alpha}{\varepsilon}\right)\right) \\
 &+ (\kappa_2 - \kappa_1) \left(\frac{1}{2} - Q\left(\frac{\alpha - \theta_{th}}{\varepsilon}\right)\right) \\
 &+ Q\left(\frac{\beta_{th} - \alpha}{\varepsilon}\right) G(\sqrt{A}, \alpha) \\
 &= \frac{1}{2} \left(\kappa_1 + \kappa_2 + Q\left(\frac{\alpha}{\varepsilon}\right)\right) + \left(\frac{1}{2} - \kappa_1\right) Q\left(\frac{\alpha - \gamma_{th}}{\varepsilon}\right) \\
 &- \kappa_1 Q\left(\frac{\beta_{th} - \alpha}{\varepsilon}\right) + (\kappa_1 - \kappa_2) Q\left(\frac{\alpha - \theta_{th}}{\varepsilon}\right) \\
 &+ \left(Q\left(\frac{\beta_{th} - \alpha}{\varepsilon}\right) - Q\left(\frac{\alpha + \beta_{th}}{\varepsilon}\right)\right) G(\sqrt{A}, \alpha). \quad (22)
 \end{aligned}$$

The inequality in the above expression is obtained similarly to the second inequality in (21). Specifically, we employ the fact that although  $P_e^{\hat{\alpha}} < \kappa_1$  when  $\hat{\alpha} > \gamma_{th}$ , the bound is tighter, that is,  $P_e^{\hat{\alpha}} < \kappa_2 < \kappa_1$ , when  $\alpha > \hat{\alpha} > \theta_{th}$ . For a given  $\varepsilon$ , if the final expression in (22) is less than or equal to  $G(\sqrt{A}, \alpha)$ , then  $P_e < G(\sqrt{A}, \alpha)$  is obtained. Therefore, under the conditions in the proposition, if the inequality in (16) is satisfied for a given  $\varepsilon$ , it is sufficient to conclude that stochastic signaling performs better than conventional signaling.  $\square$